

# Cost functions for mainline train operations and their application to timetable optimization

Aris Pavlides  
Centre for Transport Studies  
University College London  
Gower Street, London WC1E 6BT  
United Kingdom

Andy H. F. Chow  
(corresponding author)  
Centre for Transport Studies  
University College London  
Gower Street, London WC1E 6BT  
United Kingdom  
[andy.chow@ucl.ac.uk](mailto:andy.chow@ucl.ac.uk)  
Tel: +44 207 679 2315, Fax: +44 207 679 3042

5168 words + 7 figures = 6918 words

August 1, 2015

1 **ABSTRACT**

2 This paper discusses a set of cost functions for timetabling mainline train services. Mainline train  
3 services are generally heterogeneous which consist of passenger and freight trains, slow and ex-  
4 press services, domestic and international connections, etc. The feasibility of a timetable is subject  
5 to a number of factors including availability of trains and crew, infrastructure capacity, and travel  
6 demand. With the complex nature of modern railway systems and the heterogeneity of rail traffic,  
7 deriving satisfactory train service schedules for passengers, train operators, and infrastructure man-  
8 ager is always a challenge. The cost functions presented here are used as indicators for evaluating  
9 different performance associated with the corresponding timetable. The performances of interest  
10 include carbon, capacity, cost, and customer satisfaction. These four performance indicators are  
11 also identified as the '4C' criteria by the railway industry in Great Britain. These 4C criteria are  
12 set in order to address the need to improve customer satisfaction (e.g. by providing more punctual  
13 service) and operational capacity, while decreasing operational cost and carbon emission. We will  
14 also demonstrate the application of these cost functions into an optimization framework which  
15 derives optimal timetable for heterogeneous train services. The method is applied to Brighton  
16 Main Line in south-east England as a case study. The results reveal that overall performance of  
17 the railway systems can be achieved by re-scheduling and re-sequencing the train services through  
18 the optimization framework, while this may have to come at the expense of slow and local train  
19 services if the optimization is not properly formulated.

20  
21 **Keywords:** train scheduling, capacity, punctuality, multi-objective optimization, genetic algorithm

## 22 INTRODUCTION

23 Railways are generally considered to be sustainable and green compared with other modes of  
24 transport. The significance of railway systems can be reflected by the amount of investment made  
25 around the globe, exemplified by the number of high speed railways (HSR) projects, in recent  
26 years. In the UK, we have seen recently a number of large investment programmes including  
27 the Crossrail project, focusing on improving reliability, journey times, and capacity of the London  
28 transport network (1). In Hong Kong, over the next decade the MTR Corporation will complete five  
29 new strategic rail extensions in Hong Kong and mainland China (2). Nevertheless, this sustained  
30 demand has been placing tremendous pressure on the railway infrastructures. Due to the tight  
31 fiscal, physical and environmental constraints, continuous construction of new tracks and purchase  
32 of rolling stocks will not be a sustainable solution. Consequently, we will have to rely on effective  
33 utilization of existing infrastructures in terms of timetabling the train services.

34 This study looks at the issue of timetabling mainline train services for improving the overall  
35 efficiency of the rail systems. Mainline train services generally refer to connections between cities  
36 as opposed to the local metro services, while timetabling is regarded as the process of deriving a  
37 feasible schedule for a given set of train lines over a specific route through specifying the associated  
38 arrival and departure times at each designated point. Mainline services are generally heterogeneous  
39 which consist of passenger and freight trains, slow and express services, domestic and international  
40 connections, etc. Moreover, the feasibility of a timetable is also subject to a number of external  
41 factors including availability of trains and crew, infrastructure capacity, and travel demand. As a  
42 consequent, deriving a satisfactory timetable for different stakeholders including passengers, train  
43 operators, and infrastructure manager is always a challenge.

44 In this paper, we address the mainline train timetabling problem by using an optimization  
45 approach. A prerequisite for formulating the optimization problem is to define a set of cost (or  
46 objective) functions that can reflect different performances of interest to different stakeholders.  
47 Following (3), we identify Carbon, Capacity, Cost, and Customer satisfaction as the four main  
48 aspects of interest. Chen and Roberts (4) and Roberts et al. (5) further categorize and discuss them  
49 according to the associated relevance to different stakeholders. These four aspects are regarded as  
50 the '4C' criteria by the railway industry in UK. The 4C criteria are set in order to address the need to  
51 improve customer satisfaction (e.g. by providing more punctual service) and operational capacity,  
52 while decreasing operational cost and carbon emission. We believe these four are also among the  
53 main objectives in railway sector in other countries apart from the UK. A set of cost functions is  
54 formulated to reflect the performance of each timetable in terms of the '4C'. The cost functions  
55 are then incorporated in a multi-objective optimization framework (see e.g. (6, 7, 8, 9, 10)) for  
56 deriving an optimal timetable following the setting of the cost functions or objectives.

57 The optimization framework is applied to the Brighton Main Line (BML) in south-east  
58 England as a case study. It is noted that different cost functions have different dimensions. This  
59 study adopts the monetary values suggested by the Department for Transport (11) in the UK to  
60 convert and integrate all costs into monetary units. However, the proposed approach is generic  
61 and will be applicable to different systems by revising the conversion factors according to different  
62 operators' or countries' needs. The results obtained from BML reveal that overall performance of  
63 the railway systems can be achieved by re-scheduling and re-sequencing the train services through  
64 the optimization framework, while this may have to come at the expense of slow and local train  
65 services if the optimization is not properly formulated.

66 The rest of the paper is organized as follows: the next section starts with introducing the

67 specification of timetable in an optimization framework and its associated operational constraints.  
 68 It is then followed by discussion of different performance indicators related to 4C and formulation  
 69 of the associated cost functions. The cost functions are then used to formulate a multi-objective  
 70 optimization problem for train timetabling. We also discuss the complexity of the timetabling  
 71 problem and present a genetic algorithm (GA) based solution approach. The optimization frame-  
 72 work is applied to a case study of Brighton Main Line which is used to demonstrate the proposed  
 73 method and the results are discussed. Finally, the paper concludes with some final remarks and  
 74 suggestion for future work.

75

## 76 SPECIFICATION OF TIMETABLE AND ASSOCIATED CONSTRAINTS

A timetable is typically incorporated through specifying the arrival  $\tau_{n,s}$  and departure times  $\sigma_{n,s}$  of each train  $n$  over a set of control points  $s$  (which can be a station, junction, etc.) along its service route. An example is shown in Figure 1 in which the horizontal and vertical axes represent the time and position along the train route respectively. Each line on the diagram represents a train run which is specified by a series of departure  $\sigma_{n,s}$  and arrival times  $\tau_{n,s}$  at station  $s$  for each train  $n$  as specified by the timetable. Given a set of  $\sigma_{n,s}$  and  $\tau_{n,s}$ , we can derive the running time  $T_{n,s}$  of each train  $n$  between station  $s$  and  $s + 1$  as

$$T_{n,s} = \tau_{n,s+1} - \sigma_{n,s}, \quad (1)$$

and also the dwell time  $D_{n,s}$  of train  $n$  at station  $s$

$$D_{n,s} = \sigma_{n,s} - \tau_{n,s}, \quad (2)$$

The setting of the variables  $\sigma_{n,s}$  and  $\tau_{n,s}$  will be subject to a set of operational constraints in practice. We first have the minimum section running time constraints to reflect the speed limit imposed on each track section  $(s, s + 1)$ :

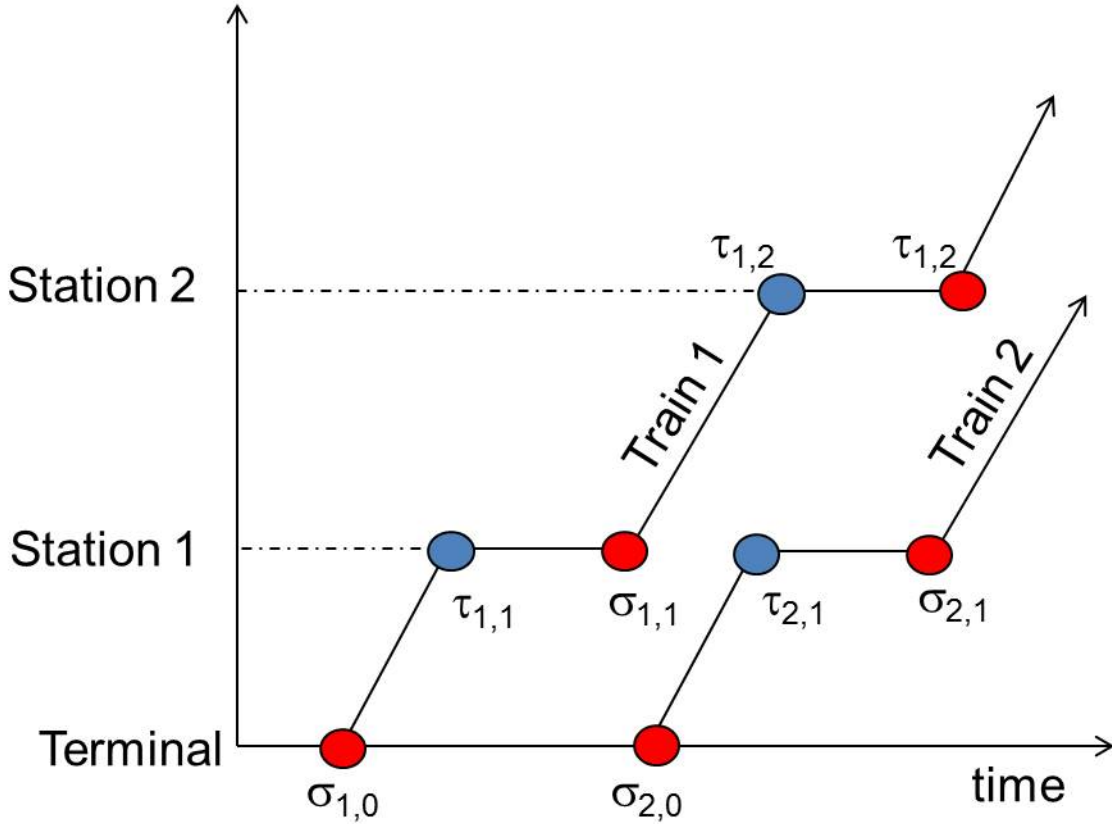
$$\tau_{n,s+1} \geq \sigma_{n,s} + \frac{\Delta_{s,s+1}}{v_n^*}, \quad (3)$$

where  $\Delta_{s,s+1}$  is the distance between stations  $s$  and  $s + 1$ ,  $v_n^*$  is the maximum speed limit for train  $n$  traveling from station  $s$  toward  $s + 1$ . Moreover, we also have the minimum dwell time constraints which define the minimum time have to be spent by each train  $n$  at station  $s$ :

$$\sigma_{n,s} - \tau_{n,s} \geq d_{n,s}^*, \quad (4)$$

77 The minimum dwell time  $d_{n,s}^*$  imposed here will typically be determined by a number of factors  
 78 on the demand side such as demand level of passengers or freight for that specific train at that  
 79 specific station, and/or the consideration of connectivity where it is necessary to ensure a long  
 80 enough dwell time for passengers or goods to transfer from one train to another at the station or  
 81 interchange (12).

Finally, to implement the signaling system, each track section is further disaggregated into a series of blocks. Under the current fixed block signaling systems in practice, each block can only accommodate up to one train at a time to ensure safe operations (see Figure 2). Referring to Figure 2, denote the arrival and departure times of train  $n$  at block  $j$  between station pair  $(s, s + 1)$



**FIGURE 1 A realization of timetable with train diagram**

as  $\sigma_{n,s,j}$  and  $\tau_{n,s,j}$  respectively. The shaded region in the figure represents the location and time period (during times  $t_{in}$  and  $t_{out}$ ) that is occupied by the train of interest during which other trains are prohibited from entering. Following the specification in the current UIC (International Union of Railways) operational code (13), we have

$$t_{in} = \tau_{n,s,j} + \frac{\delta_{n,j}}{v_{n,s,j}}, \quad (5)$$

where  $\delta_{n,j}$  is the visual distance of train  $n$  to the entrance of block  $j$ ;  $v_{n,s,j}$  is the nominal speed of train  $n$  traveling through block  $j$ . The time  $t_{in}$  represents the time when the driver of train  $n$  observes the signal aspect at block  $j$  and starts to take action(s) accordingly. Moreover,

$$t_{out} = \sigma_{n,s,j} + \frac{L_n}{v_{n,s,j}}, \quad (6)$$

where  $L_n$  is the length of train  $n$ . The time  $t_{out}$  represents the time when the tail of the train  $n$  clears from the block section. Because of the signaling system, it is expected congestion will occur when the train volume on a track section is high (14, 15). Following (5) and (6), the signal blocking constraint can then be written mathematically for all station pairs  $(s, s + 1)$  and signal blocks  $j$  as

$$\tau_{n+1,s,j} \geq \sigma_{n,s,j} + \frac{L_n}{v_{n,s,j}}, \quad (7)$$

82 in which train  $n + 1$  is the train following immediately after train  $n$ .

83

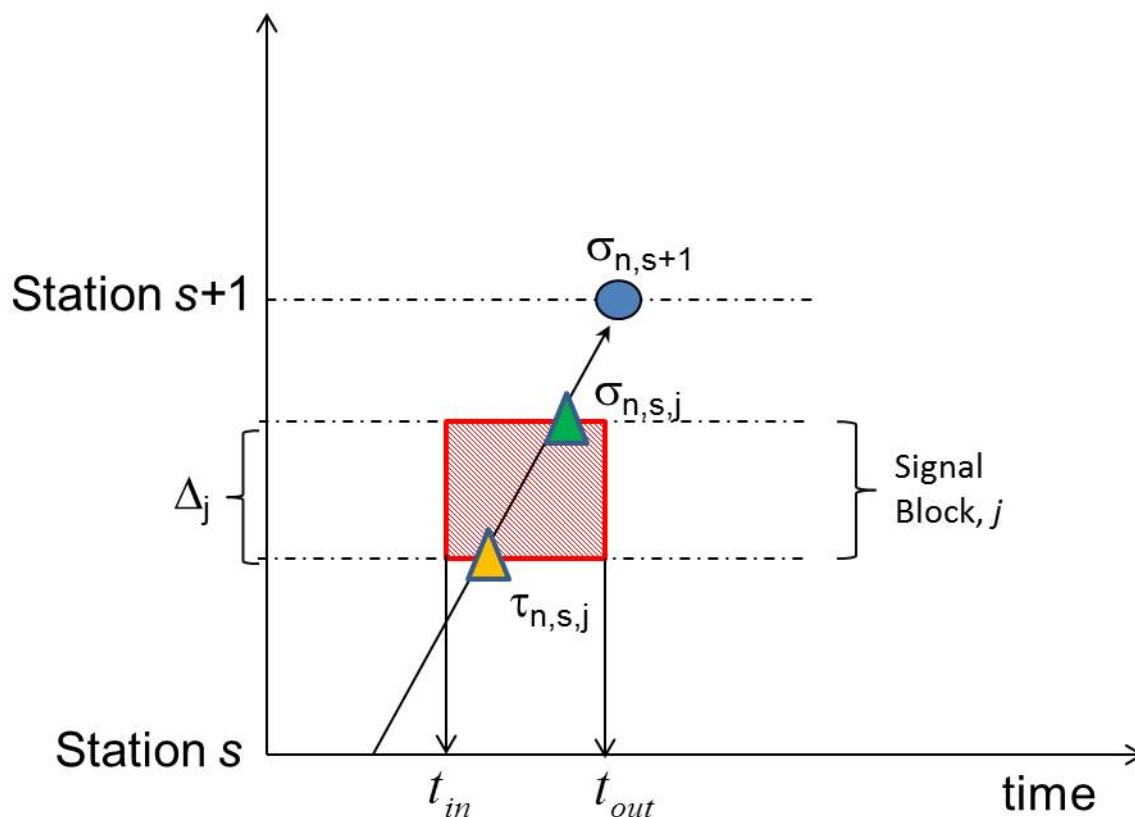


FIGURE 2 Representation of fixed block system

#### 84 PERFORMANCE INDICATORS AND COST FUNCTIONS

85 With the timetable and the associated constraints specified, we can then formulate the cost func-  
 86 tions to be used in the optimization framework which reflect various performance in railway  
 87 timetabling and operations. Following the comprehensive review in (4) and (5), we have selected  
 88 five representative performance indicators in the railway industry: train running times, customer  
 89 waiting times, service punctuality, utilization of trains and track resources. It is recognized that  
 90 the first three performances will be specifically interesting to customers (passengers and freight  
 91 companies), utilization of trains will be interesting to Train Operators, and utilization of track will  
 92 be interesting to Infrastructure Manager.

93

#### 94 Running times of trains

Running times  $T_{n,s}$  of trains  $n$  over all section  $(s, s + 1)$  can be obtained from Equation (1) in the  
 previous section following the specification of timetable variables  $\sigma_{n,s}$  and  $\tau_{n,s}$  as discussed. Given

all running times  $T_{n,s}$ , we define the cost associated with the running time components as

$$C_T = \hat{c}_T \sum_{n=1}^N \sum_{s=1}^S T_{n,s} p_{n,s}, \quad (8)$$

95 where  $N$  and  $S$  represent the total number of trains and stations in the system respectively. The  
 96 variable  $p_{n,s}$  is a quantity associated with the demand which represents the number of passengers  
 97 (or amount of goods) on train  $n$  running between stations  $s$  and  $s + 1$ . Determining this  $p_{n,s}$  will  
 98 require detailed origin-destination (OD) survey which can be difficult in practice. The quantity  
 99  $p_{n,s}$  may be dropped from (8) if such OD information is not available, and this will result in the  
 100 optimizer treating each train  $n$  equally. With this  $p_{n,s}$ , the corresponding timetable will then give  
 101 higher priority to trains carrying more passengers or goods after optimization. Finally, the notation  
 102  $\hat{c}_T$  represents a monetary cost associated with running times, where some examples can be found  
 103 in (11, 16, 17). We will have further discussion on the choice of this  $c_T$  and other monetary cost  
 104 coefficients in latter section.

105

### 106 **Waiting times of passengers (or goods)**

Estimating the cost associated with waiting times first requires knowledge of  $\lambda_s(t)$  which denotes the profile of demand for service at station  $s$  over time  $t$ . Fundamental queueing analysis (e.g. (18)) gives the total waiting time  $W$  (in the unit of [persons-time] or [goods-time]) as

$$W = \sum_{s=1}^S \sum_{n=1}^{N_s-1} \int_{\tau_{n,s}}^{\tau_{n+1,s}} \lambda_s(t) dt^2, \quad (9)$$

where  $N_s$  is the total number of trains serving station  $s$  over the study time period. The time interval between  $\tau_{n,s}$  and  $\tau_{n+1,s}$  specify the headway of train service at station  $s$ . Equation (9) can be simplified by assuming a uniform demand  $\bar{\lambda}_s = \lambda_s(t)$  for all times  $t$  during the study period as:

$$W = \sum_{s=1}^S \sum_{n=1}^{N_s-1} \bar{\lambda}_s [\tau_{n+1,s} - \tau_{n,s}]^2. \quad (10)$$

107 As reflected from (10), the total waiting time grows linearly with the average demand rate  $\bar{\lambda}_s$  but  
 108 quadratically as the service headway increases (i.e. frequency of service decreases). However, the  
 109 uniform demand assumption made in deriving (10) may be valid for high frequency service (e.g.  
 110 metro) while it may not be appropriate for low frequency mainline services as it is known that the  
 111 arrival of passengers will cluster around the publicized scheduled service times in the timetable.  
 112 Hence some detailed survey will be needed for obtaining the demand pattern if one wants to have  
 113 a reasonable estimate of waiting times when deriving mainline timetable.

Finally, following the calculation of  $W$ , the eventual cost associated with waiting times is determined as

$$C_W = \hat{c}_W W, \quad (11)$$

114 where  $\hat{c}_W$  is the monetary cost associated with waiting times. The purpose of incorporating the  
 115 waiting time into the optimization framework is to ensure that there are enough services for num-  
 116 ber of passengers or goods at the station without creating excessive waiting times. Empirical

117 studies conducted by the UK Department for Transport (e.g. (11, 16, 17)) suggest that this  $\hat{c}_W$  will  
 118 be around two or three times larger than  $\hat{c}_T$  as the waiting time is generally regarded as a dead time.  
 119

### 120 Punctuality of service

Punctuality is measured herein as the time discrepancy between the scheduled and the actual arrival times of the train services. To quantify the punctuality in monetary unit (see (19), (20)), we adopt a schedule cost function as shown in Figure 3. In the figure,  $\tau^*$  denotes the ideal arrival time of the train service while  $\Phi$  is a time allowance for lateness (e.g.  $\Phi$  is considered to be three minutes under the UK railway operational regulations (20)). If the corresponding train is delayed by more than  $\Phi$  from the ideal arrival time  $\tau^*$ , a schedule delay cost will be imposed on the Train Operator by the Infrastructure Manager for lateness. It is considered here that this schedule delay cost increases linearly with a slope of  $\hat{c}_P$  over arrival time  $\tau$ , where  $\tau \geq \tau^* + \Phi$ . This penalty rate  $\hat{c}_P$  represents the value of lost time of customers (passengers or freight companies) per unit lateness in time (20, 21). Following this linear specification, the total schedule delay cost associated with punctuality can be determined, taking the arrival of passengers and/or goods into account, as

$$C_P = \hat{c}_P \sum_{s=1}^S \sum_{n=1}^{N_s-1} \int_{\tau_{n,s}}^{\tau_{n+1,s}^*} \lambda_s(t) (\tau_{n+1,s} - \tau_{n+1,s}^* - \Phi)^+ dt, \quad (12)$$

where  $\tau_{n+1,s}^*$  is the ideal arrival time for train  $n + 1$  at station  $s$ ,  $(\tau_{n+1,s} - \tau_{n+1,s}^* - \Phi)^+ = \max[(\tau_{n+1,s} - \tau_{n+1,s}^* - \Phi), 0]$ . Similar to (10), Equation (12) can be simplified by assuming uniform arrival  $\bar{\lambda}_s = \lambda_s(t)$  for all times  $t$  as

$$C_P = \hat{c}_P \sum_{s=1}^S \sum_{n=1}^{N_s-1} \bar{\lambda}_s (\tau_{n+1,s}^* - \tau_{n,s}) (\tau_{n+1,s} - \tau_{n+1,s}^* - \Phi)^+. \quad (13)$$

121 Finally, it is noted that this punctuality cost analysis is generally applicable to other schedule  
 122 cost functions, apart from the linear assumption in Figure 3, by revising the cost function term  
 123  $(\tau_{n+1,s} - \tau_{n+1,s}^* - \Phi)^+$  in (12) and (13) accordingly.  
 124

### 125 Utilization of trains

If the on-board loading  $p_{n,s}$  of each train  $n$  between each station pair  $(s, s + 1)$  is available, we can also derive a cost associated with the utilization of trains as

$$C_L = \hat{c}_L \sum_{n=1}^N \sum_{s=1}^S \left(1 - \frac{p_{n,s}}{p_n^*}\right), \quad (14)$$

126 where  $p_n^*$  is the physical holding capacity of train  $n$  for passengers or goods,  $\hat{c}_L$  is the monetary  
 127 cost associated with per unit lost due to inefficient use of train holding capacity. The cost  $C_L$  will  
 128 be an useful component to be included from Train Operators' perspective for deriving effective  
 129 strategies transporting passengers and goods with the least number of trains.  
 130



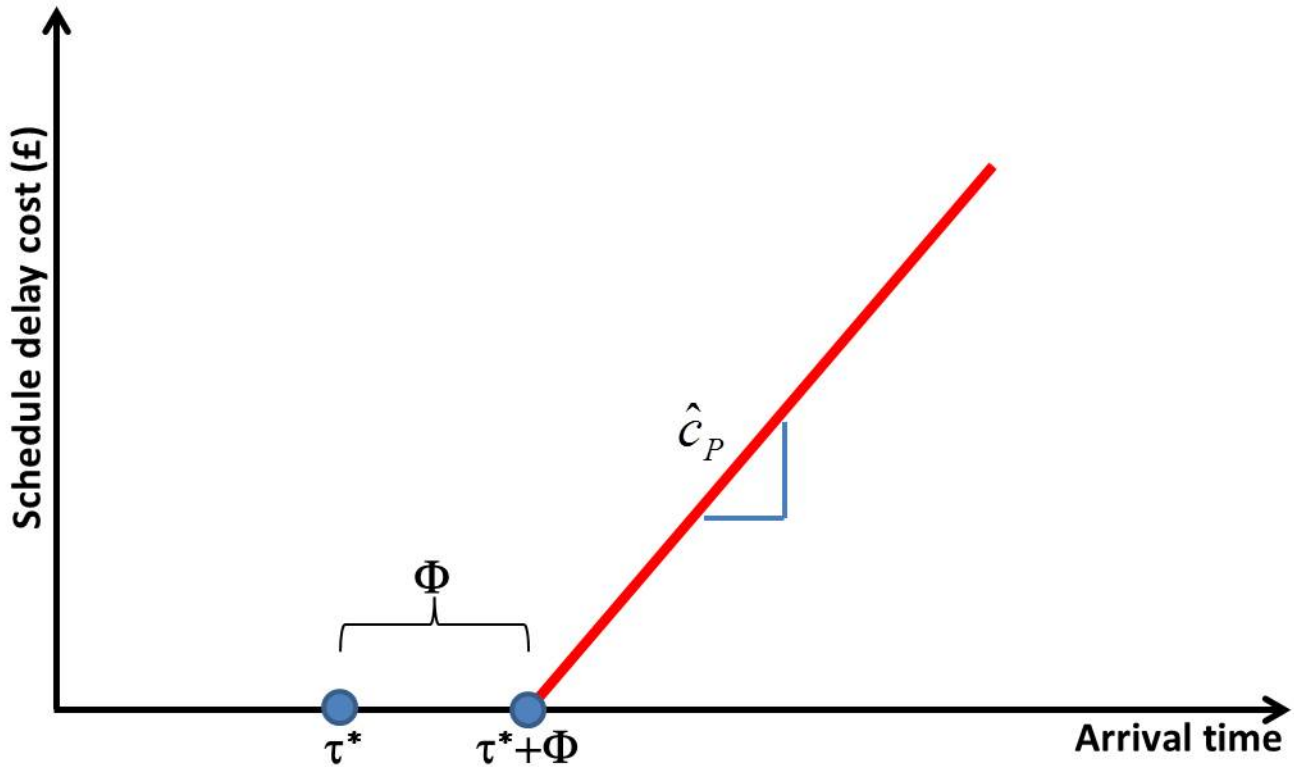


FIGURE 3 Schedule delay cost function

131 **Utilization of track capacity**

Utilization of track capacity is measured by using the occupation measure specified in the UIC 406 'Capacity' code (13). The occupation of a block section is defined as the total time that the block is occupied by trains within a specific study period, divided by the length of the study period. Referring to an example in Figure 4 which shows two trains passing through a block section within a specific time period  $T$ . Denote  $(t_{in})_{n,j}$  and  $(t_{out})_{n,j}$  respectively the entry and exit times of train  $n$  to the block ( $j$ ) of interest. The occupation ratio for this block  $j$  is calculated as

$$occ_j = \frac{\sum_{n=1}^{N_j} [(t_{out})_{n,j} - (t_{in})_{n,j}]}{T}, \quad (15)$$

where  $N_j$  is the total number of trains passing the block in  $T$ . We can then come up with a network-wide measure of track utilization as over all track sections between stations  $s$  and  $s + 1$  in the system as:

$$OCC = \sum_{s=1}^S \sum_{j=1}^{J_s} occ_j, \quad (16)$$

where  $J_s$  is the number of blocks along track section between stations  $s$  and  $s + 1$ . Different from previous UIC 405 standard (22) which only considers only the number of trains passing the

block sections, the UIC 406 approach captures the heterogeneity and speed differentials among trains through considering the time occupied by trains. Following the  $OCC$  calculated by (15), the monetary cost associated with track utilization is determined as:

$$C_U = \hat{c}_U(1 - OCC), \quad (17)$$

132 where  $\hat{c}_U$  is recognized as the cost per unit lost of track occupation. This  $C_U$  will be an useful  
 133 indicator for Infrastructure Manager which aims to maximize the efficiency of utilizing limited  
 134 infrastructure capacity.

135

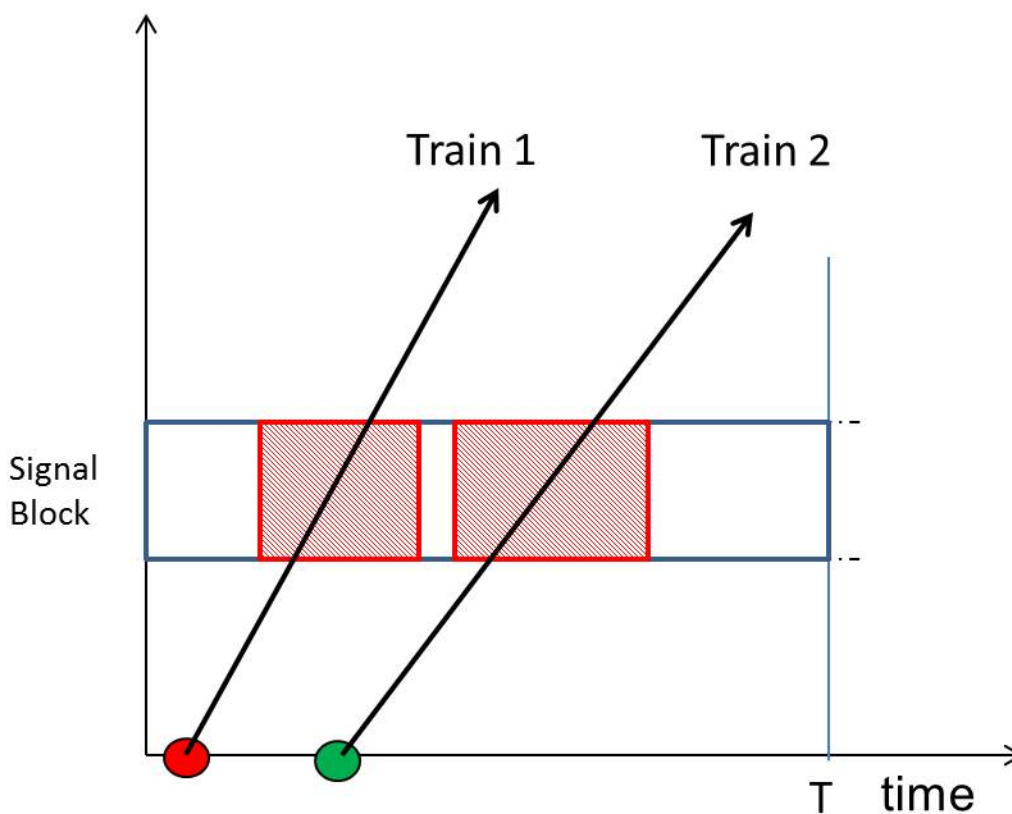


FIGURE 4 Measure of utilization of track (13)

### 136 Discussion

137 This section presents five performance indicators widely used in the railway industry and the for-  
 138 mulations of the corresponding cost functions. It should be emphasized that the five performance  
 139 indicators considered herein are mainly for illustration purposes and readers can incorporate other  
 140 performances of interest such as energy consumption and connectivity in the proposed optimization  
 141 framework. Detailed discussions of other performance indicators and the associated cost functions  
 142 can be found in (4) and (5).

143 It is understood that different stakeholders have different level of interest in different per-  
 144 formances. For example, customers will obviously be concerned about the running time, waiting

145 time, and punctuality of their services, while they will be less interested in how the track or rolling  
 146 stock resources are utilized. Infrastructure manager and train operators on the other hand will be  
 147 very careful about planning the use of their resources (track and trains) in order to come up with the  
 148 most cost-effective operational strategy. Unfortunately, different performances are often in conflict  
 149 of each other. Utilization of track can be enhanced by running more trains within a given time pe-  
 150 riod, while this could have an adverse effect on punctuality as it will be more likely generating  
 151 delays due to congestion. The conflict between different performance indicators and stakeholders'  
 152 interest calls for the use of multi-objective optimization technique to come up with a timetable that  
 153 can maximize the overall performance of the system while taking the conflicts into account.

## 155 APPLICATION TO TIMETABLE OPTIMIZATION

The cost functions developed in previous section are applied to formulate a multi-objective opti-  
 mization problem. The optimization aims to determine the train timetable, in terms of arrival  $\tau_{n,s}$   
 and departure times  $\sigma_{n,s}$  for all trains  $n$  over all stations  $s$ , such that the following total cost is  
 minimized:

$$C = C_T + C_W + C_P + C_L + C_U. \quad (18)$$

156 The cost in (18) is in monetary unit and its cost components are integrated through the monetary  
 157 cost coefficients:  $\hat{c}_T, \hat{c}_W, \hat{c}_P, \hat{c}_L$ , and  $\hat{c}_U$  as discussed. The cost minimization problem is subject to  
 158 the operational constraints (3), (4), and (7).

159 The train timetable optimization problem is combinatoric that involves different feasible  
 160 combinations of  $\tau_{n,s}$  and  $\sigma_{n,s}$  representing different sequencing and scheduling of trains (23, 24).  
 161 Considering a scenario where there are  $N$  trains to schedule, the number of possible sequences for  
 162 scheduling these trains will be  $N!$ . This has not included the infinite number of ways of setting the  
 163 departure and arrival times of these trains along the service route given a sequence.

164 To derive a solution within a reasonable time, an optimal sequence and times of departures  
 165 of trains from their terminals is searched by using a genetic algorithm (GA). The genetic algorithm  
 166 starts with a population (e.g. with a size of around 100) of randomly generated sequences of trains  
 167 which are regarded as 'chromosomes'. Each chromosome is a combination of binary (0-1) bit  
 168 representing different train sequences. For example, consider there are three trains (A, B, C) with  
 169 different service paths and characteristics to schedule. This gives a total of  $3! = 6$  possible se-  
 170 quences: ABC, ACB, BAC, BCA, CAB, and CBA. This can be represented by a set of 3-bit binary  
 171 chromosomes (which gives a total of possible  $8 (=2^3)$  combinations). Given the train sequence,  
 172 the corresponding departure  $\sigma_{n,s}$  and arrival times  $\tau_{n,s}$  of each train is then computed by using a  
 173 greedy search approach in the second stage. The greedy search strategy determines the  $\sigma_{n,s}$  and  
 174  $\tau_{n,s}$  as the earliest times that each train can proceed subject to constraints (3), (4), and (7). In case  
 175 of a conflict occurs when two (or more) trains meet at a junction along their service lines, priority  
 176 is given based upon the first-come-first-serve principle.

With the set of chromosomes containing information of sequence and departures of trains,  
 the optimizer starts with the 'reproduction' step which reproduces chromosomes according to their  
 'fitness' values in the next iteration. The fitness value is calculated based upon the value of total  
 cost (18) associated with the train sequence and departures specified in the chromosome. Essen-  
 tially a higher fitness value will be assigned to a chromosome if the chromosome achieves lower

total cost, and the fitness function  $FIT_i$  for each chromosome  $i$  is defined as:

$$FIT_i = \frac{A_i}{\sum A_i}, \quad (19)$$

where

$$A_i = \exp\left(\left(\frac{C_{max} - C_i}{C_{max} - C_{min}}\right)p\right), \quad (20)$$

177 in which  $C_i$  is the value of total cost calculated from (18) based on the sequence and departures of  
 178 trains specified in chromosome  $i$ ,  $C_{max}$  and  $C_{min}$  are respectively the maximum and minimum cost  
 179 values identified in the current iteration of optimization,  $p$  is a parameter tuning the fitness function  
 180 for maximizing the efficiency of the optimizer where it is set to be 5 here. The chromosomes are  
 181 'reproduced' in proportion to their fitness value  $FIT_i$  calculated above.

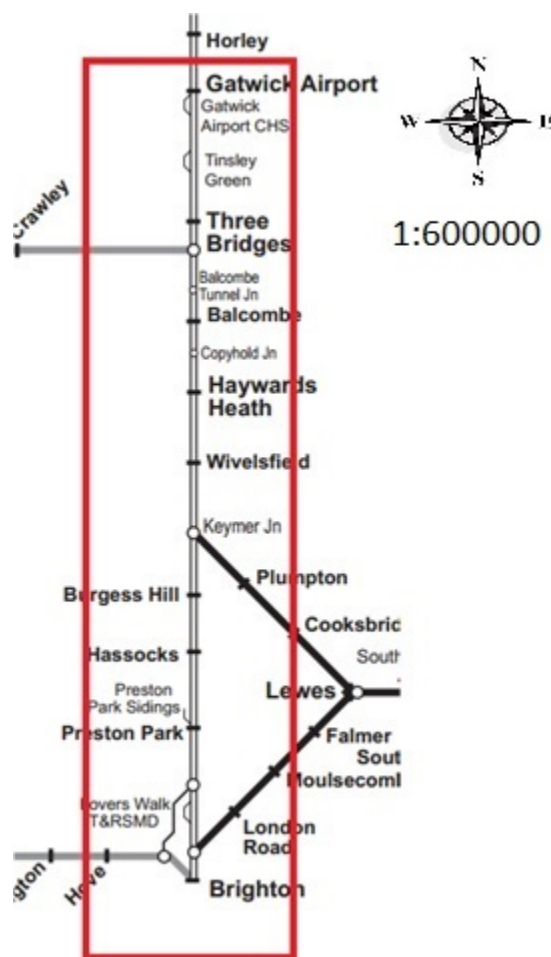
182 Following the reproduction step, the 'crossover' operation will then randomly select and  
 183 'mate' two chromosomes (regarded as 'parents'). The GA optimizer separates each 'parent' chro-  
 184 mosome into two parts, swaps with each other, and forms a new pair of chromosomes (which are  
 185 regarded as 'children'). This crossover process is for generating the next set of population with  
 186 some entirely new characteristics with respect to the previous population and hence avoiding the  
 187 optimisation process from trapping into local optima. Finally, the 'mutation' process randomly se-  
 188 lects some bits in the population with a predefined probability (typically 0.005 - 0.01) and 'mutate'  
 189 (i.e. a '0' bit will be changed to '1' arbitrarily, and a '1' bit will be changed to '0'). This is again  
 190 to prevent the optimization process from trapping into local optima. The GA optimization process  
 191 above (reproduction-crossover-mutation) will continue until the predefined maximum number of  
 192 iterations (e.g. 20 - 30) is reached. Further details of GA can be found in a number of literature  
 193 including (25).

194

## 195 **CASE STUDY - BRIGHTON MAIN LINE (UK)**

196 The optimization framework is applied to the Brighton Main Line in southeast England (Figure  
 197 5). The Brighton Main Line is approximately 80-km long electrified connection linking London  
 198 Victoria and London Bridge with Brighton via East Croydon and Gatwick Airport. The line itself  
 199 has a complex structure with a variable number of tracks (four tracks from London down to Bal-  
 200 combe Tunnel Junction and two tracks thereafter), different speed limits along the line, multiple  
 201 branch lines (e.g. at Junctions Horsham, Lewes), and sidings (e.g. along Ardingly, Lovers Depot).  
 202 Passenger operators that operate on the BML include Southern and First Capital Connect. We  
 203 select the section between Gatwick Airport and Brighton which is highlighted in Figure 5. This  
 204 is one of the busiest sections along BML. The study period is 08:00 - 10:00, which is regarded  
 205 as the morning peak, on weekdays. During the study period there is currently a total of 22 trains  
 206 running from Brighton toward Gatwick and hence Central London (the 'Up' direction) and 18  
 207 trains running from Gatwick toward Brighton (the 'Down' direction). We derive this 'base case'  
 208 train timetable with information obtained from Network Rail. The idea is to derive an optimized  
 209 timetable from the proposed optimization framework with the same number of trains within the  
 210 same study period. We then compare the 'optimized' timetable with this 'base case' timetable to  
 211 see how much improvement, in terms of reduction in costs, can be achieved in different aspects  
 212 through re-sequencing and re-scheduling. There are two different train classes running through

213 the section during the study period: Classes 375 and 442 with Class 375 used for the express  
 214 connection. Both train classes are used for passenger transport, while it should be noted the pro-  
 215 posed optimization framework presented in this paper can capture any number and type of train  
 216 classes including freight train. Finally, it is noted that the actual origin-destination demand matrix  
 217 is not made available so that an average demand rate ( $\bar{\lambda}_s$ ) and train loading ( $p_{n,s}$ ) will have to be  
 218 estimated from field observations on a weekday. In general it is recognized that the demands at  
 219 the major stations including Gatwick Airport, Three Bridges and Brighton are higher than other  
 220 stations which is expected as these are some major hubs or interchanges along the line



**FIGURE 5 Test network - Brighton Main Line (UK)**

221 The coefficients in the cost function (18) are set here with official documents by the British  
 222 government organizations. On the customer side, the monetary cost  $\hat{c}_T$  associated with running  
 223 times of train is set to be £5.76 (per person-hour), while the monetary costs  $\hat{c}_W$  and  $\hat{c}_P$  are both  
 224 £14.4 (per person-hour) for waiting times and punctuality respectively. The figures are set ac-  
 225 cording the 'webTAG Unit 3.5.6' guidance (11) published by UK Department for Transport which  
 226 specifies the values of time of travelers based on an empirical study conducted by University of  
 227 Leeds (26). The monetary costs of waiting times and punctuality are around two times higher  
 228 than the one for running time. It is because waiting times and delays due to lateness are generally

229 regarded as non-productive dead loss. From the perspective of Train Operators and Infrastructure  
230 Manager, the costs  $\hat{c}_U$  and  $\hat{c}_L$  respectively for utilization costs of trains and track are set to be  
231 £350 (per train) following the current track usage price published by UK Network Rail (27) which  
232 specifies the cost for deploying a train on the track.

233 Given the network configuration and cost coefficients, the total cost (18) associated with the  
234 existing timetable is determined as £66.7k. The breakdown of this total cost into its components  
235 is shown in Figure 7. It is shown that the majority ( $\sim 85\%$ ) of the cost is associated with the  
236 customer related components: running times, waiting times, and punctuality. As a consequence, it  
237 can be expected that the eventual optimized timetable would favor customers over Train Operators  
238 and Infrastructure Manager. This however can be modified with revised formulation of the cost  
239 functions and coefficients.

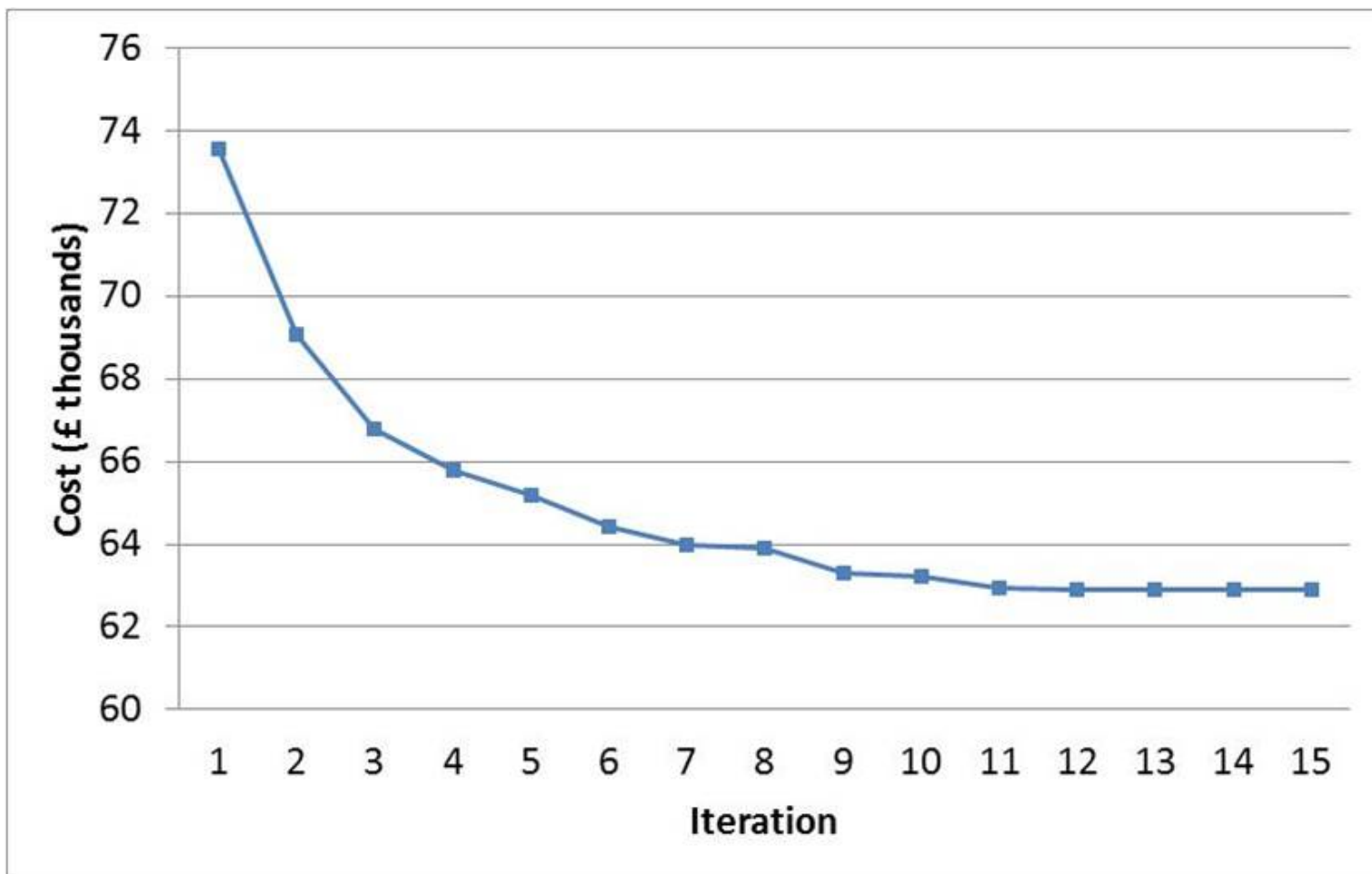
240 Figure 6 shows the progress of the optimization process in which the value of total cost  
241 is reduced gradually from the initial value £73.6k with randomly generated timetables to eventual  
242 £62.9k with the optimized one after 15 iterations given the same number of trains to schedule, the  
243 same number of passengers to serve, over the same period. The optimization process takes five  
244 minutes to complete on a standard Windows 7 (64-bit) desktop computer. Similar to other im-  
245 plementations of genetic algorithm (e.g. (25)), the most significant improvements are observed in  
246 the first few generations while the optimization process gradually converges slowly to the ultimate  
247 final solution at latter iterations.

248 Figure 7 further compares the cost components before and after the optimization. As afore-  
249 mentioned, the optimization mainly benefit the customers' costs due to their large portion in the  
250 cost components. The reduction in waiting times comes from assigning more priority to trains  
251 (e.g. the express or 'fast' trains) serving major stations with higher demand over other trains serv-  
252 ing local area. This can be revealed from Figure 8 which compares the train diagrams under the  
253 nominal and optimized timetable toward the end of the study period (after 09:00). Under the orig-  
254 inal timetable, the fast trains (Class 375) are hindered by the slow train (Class 442) highlighted in  
255 the figure. This leads to higher costs associated with running times and hence potentially waiting  
256 times and punctuality. After optimization, more slow trains are scheduled toward the end of the  
257 study period with an objective to give way to the faster Class 375 trains in the front. As the number  
258 of available trains is considered to be fixed, the improvements in punctuality as well as utilization  
259 of trains and track, are insignificant. This however can be modified by allowing more (or less)  
260 trains to be scheduled in the optimization process. One can also estimate the marginal cost of  
261 adding or reducing a train with respect to the overall system performance.

## 263 CONCLUSIONS

264 This paper presents a multi-objective optimization framework which derives optimal timetables for  
265 mainline train service that maximizes the system efficiency in various performance aspects. The  
266 performances considered herein include running times of trains, waiting times of customers for  
267 service, punctuality, utilization of trains and track. The performances considered cover different  
268 stakeholders: customers, Train Operators, and Infrastructure Manager. The contributions of this  
269 paper include specification of timetable and its associated operational constraints, formulations of  
270 cost functions reflecting the corresponding performances, and multi-objective optimization with a  
271 GA-based solution method.

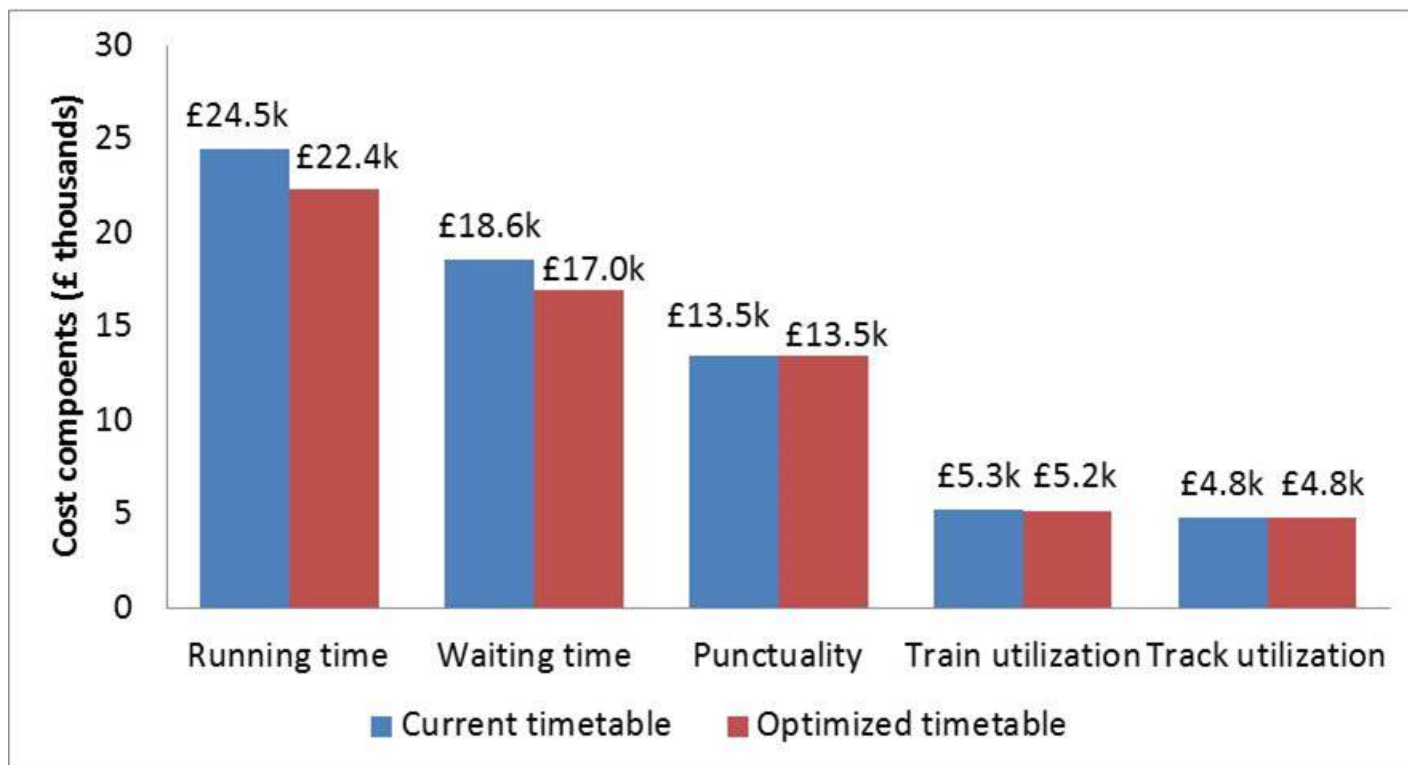
272 The optimization framework is applied to the Brighton Main Line in southeast England.



**FIGURE 6 Progression of optimization process**

273 Given the network configuration and demand, the optimal timetable is solved by a two-stage solu-  
274 tion procedure based upon genetic algorithm. The optimizer is shown to be able to reduce the cost  
275 of operations in particular in the aspects of running times, waiting time, and punctuality. Never-  
276 theless, it is revealed that this is achieved by assigning higher priority to fast express trains at the  
277 expense of slow local trains. This may not be a desirable result if one is interested in improving  
278 the equity of different service types. In particular, it is found that current policies of many Infras-  
279 tructure Managers around the world tend to favour passenger train operations over freight ones  
280 due to the higher demand for passenger trains, higher speeds, and less energy consumed. Such  
281 timetabling and capacity allocation policy however can hurt the freight train industry in the long  
282 run. Incorporating the equity of train services will be a future research direction. Finally, it is noted  
283 that the focus of the present paper lies on the formulation of cost functions and their application to  
284 timetabling instead of the optimization algorithm. We agree that it will be worthy of conducting  
285 further research on alternative algorithms for improving the quality of the optimal solutions.

286



**FIGURE 7 Cost components before and after optimization**

## 287 ACKNOWLEDGMENTS

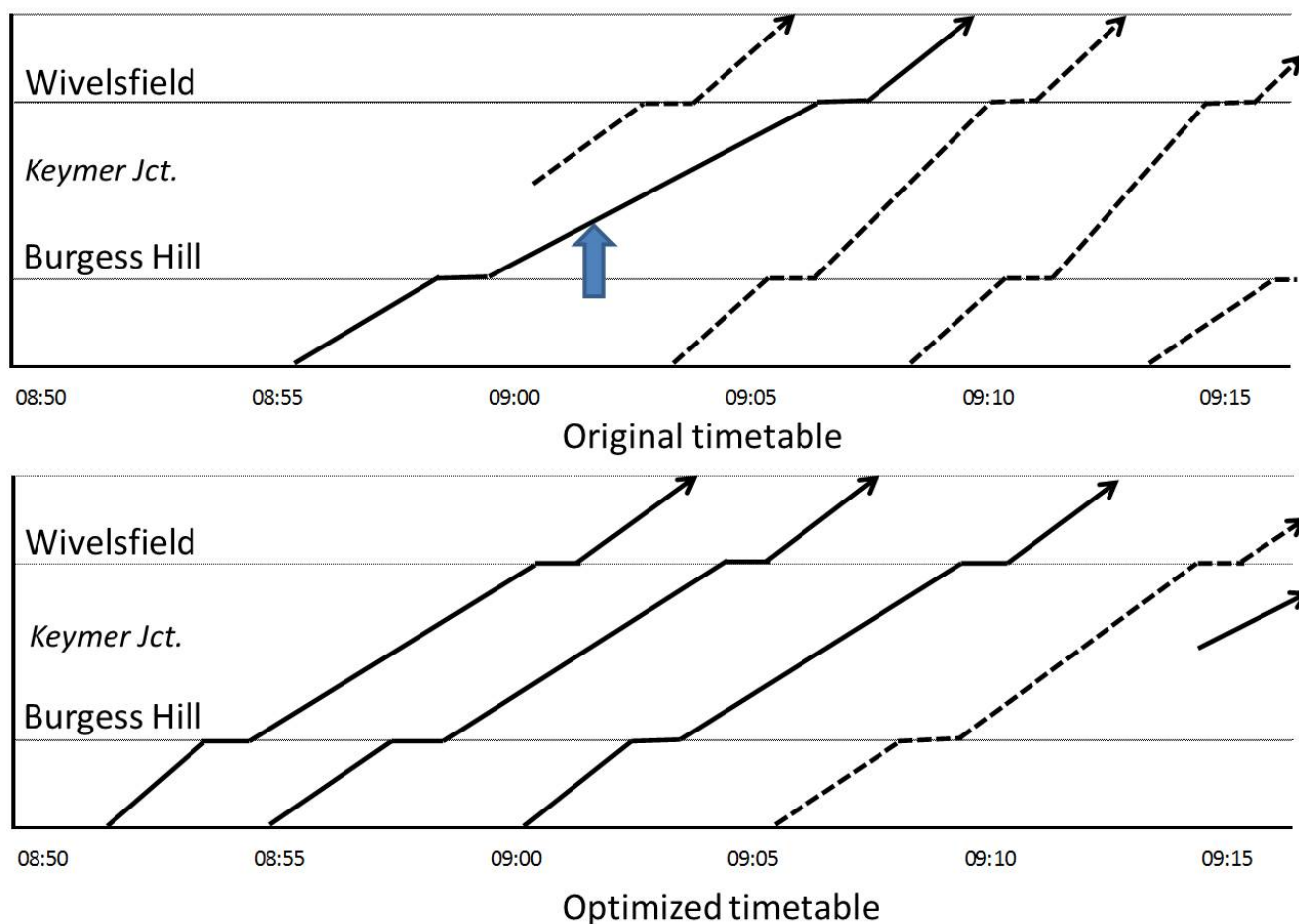
288 This study is funded by an Industrial CASE studentship awarded by UK Engineering and Physical  
 289 Sciences Research Council (EPSRC) and Rail Safety and Standard Board (RSSB). We would like  
 290 to thank Catherine Baker at RSSB for her valuable comments. The contents of this paper reflect  
 291 the views of the authors who are responsible for the facts and the accuracy of the data presented  
 292 herein. The contents do not reflect the official views or policies of RSSB, Network Rail, or any  
 293 other organization. This paper also does not constitute a standard, specification, or regulation.

294

## 295 REFERENCES

- 296 [1] UK HM Treasury, *Investing in Britain's future*, 2013.
- 297 [2] Hong Kong MTR Corporation Limited, *Sustainability Report*, 2013.
- 298 [3] UK Rail Safety and Standards Board (RSSB), *The Future Railway - The Industry's Rail Tech-*  
 299 *nical Strategy (RTS 2012)*. Technical Strategy Leadership Group (TSLG), RSSB, London,  
 300 UK, 2012.
- 301 [4] Chen, L. and Roberts, C., *Defining the Optimisation Function for Timetabling as Part of*  
 302 *Future Traffic Regulation (FuTRO)*. Report submitted to Railway Safety and Standards Board,  
 303 United Kingdom, 2012.
- 304 [5] Roberts, C., L. Chen, M. Lu, L. Dai, C. Bouch, G. Nicholson, and F. Schmid, *A Framework*





**FIGURE 8 Train diagrams before and after optimization (dotted line: fast trains; solid line: slow trains)**

305 *for Developing an Objective Function for Evaluating Work Package Solutions (Cost func-*  
 306 *tion)*. Deliverable 1.2: Optimal Networks for Train Integration Management across Europe  
 307 (ONTIME), European Commission Seventh Framework Programme (FP7), 2013.

308 [6] Ghoseiri, K., Szidarovszky, F. and Asgharpour, M. J., A multi-objective train scheduling  
 309 problem and solution. *Transportation Research Part B*, Vol. 38, No. 10, 2004, pp. 927–952.

310 [7] Bussieck, M. R., Winter, T. and Zimmermann, U. T., Discrete optimization in public rail  
 311 transport. *Mathematical Programming*, Vol. 79, No. 1-3, 2009, pp. 415–444.

312 [8] Vansteenwegen, P. and Van Oudheusden, D., Developing railway timetables which guarantee  
 313 better service. *European Journal of Operational Research*, Vol. 173, No. 1, 2006, pp. 337–  
 314 350.

315 [9] Vansteenwegen, P. and van Oudheusden, D., Decreasing the passenger waiting time for an  
 316 intercity rail network. *Transportation Research Part B*, Vol. 41, No. 4, 2007, pp. 478–492.

- 317 [10] Albrecht, T., Automated timetable design for demand-oriented service on suburban railways.  
318 *Public Transport*, Vol. 1, No. 1, 2009, pp. 5–20.
- 319 [11] UK Department for Transport, *TAG UNIT 3.5.6: Values of time and vehicle operating costs*.  
320 [www.persona.uk.com/a5dunstable/deposit-docs/DD-096.pdf](http://www.persona.uk.com/a5dunstable/deposit-docs/DD-096.pdf),  
321 2012, [Accessed on: 22 Jul 2015].
- 322 [12] Pavlides, A. and Chow, A. H. F. and Baker, C., A study of cost functions for timetabling national  
323 railway operations. *Proceedings of the 47th Annual Conference of Universities Transport Study Group*,  
324 *January 5-7. London, UK*, 2015.
- 325 [13] International Union of Railways (UIC), *UIC Code 406: Capacity*. [banportalen.banverket.se/Banportalen/upload/1753/HandbokUIC406.pdf](http://banportalen.banverket.se/Banportalen/upload/1753/HandbokUIC406.pdf), 2004,  
326 [Accessed on: 22 Jul 2015].
- 328 [14] Gibson, S., Cooper, G. and Ball, B., The evolution of capacity changes on the UK rail network.  
329 *Journal of Transport Economics and Policy*, Vol. 36, No. 2, 2002, pp. 341–354.
- 330 [15] Dinger, M. H., Lai, YC. and Barkan, C. P. L., Population Synthesis for Microsimulating  
331 Travel Behavior. *Transportation Research Record*, Vol. 2117, 2009, pp. 41–49.
- 332 [16] Wardman, M., The value of travel time - A review of British evidence. *Journal of Transport  
333 Economics and Policy*, Vol. 32, No. 3, 1998, pp. 285–316.
- 334 [17] Wardman, M., Public transport values of time. *Transport Policy*, Vol. 11, No. 1, 2004, pp.  
335 363–377.
- 336 [18] Daganzo, C. F., *Fundamentals of Transportation and Traffic Operations*. Pergamon, New  
337 York, 1997.
- 338 [19] Carey, M., Reliability of interconnected scheduled services. *European Journal of Operational  
339 Research*, Vol. 79, No. 1, 1994, pp. 51–72.
- 340 [20] UK Network Rail, *Schedule 8: Performance Regime*, 2015.
- 341 [21] UK Office of Rail Regulation, *Passenger & Freight rail performance quality report*, 2015.
- 342 [22] International Union of Railways (UIC), *UIC Code 405: Links between Railway Infrastructure  
343 Capacity and the Quality of Operations*, 1996.
- 344 [23] D’Ariano, A., D. Pacciarelli, and M. Pranzo, A branch and bound algorithm for scheduling  
345 trains in a railway network. *European Journal of Operational Research*, Vol. 183, No. 2,  
346 2007, pp. 643–657.
- 347 [24] Xu, F., B. Heydecker, A. H. F. Chow, and T. Fujiyama, Optimisation framework for rail  
348 traffic regulation and speed optimisation at a single junction. *Proceedings of International  
349 Conference of RailTokyo*, 2015, March 23-26. Tokyo, Japan.
- 350 [25] Lo, H. and A. H. F. Chow, Control strategies for over-saturated traffic. *ASCE Journal of  
351 Transportation Engineering*, Vol. 130, No. 4, 2004, pp. 466–478.

- 352 [26] Mackie, P. J., M. Wardman, A. S. Fowkes, G. Whelan, J. Nellthorp, and J. Bates, *Values of*  
353 *travel time savings in the UK*, 2003.
- 354 [27] UK Network Rail, *Track Usage Price List, 2015/16 prices*. [www.networkrail.co.uk/](http://www.networkrail.co.uk/using-our-network/cp5-access-charges/)  
355 [using-our-network/cp5-access-charges/](http://www.networkrail.co.uk/using-our-network/cp5-access-charges/), 2014, [Accessed on: 22 Jul 2015].